11099CH07

अध्याय 7

सहसंबंध

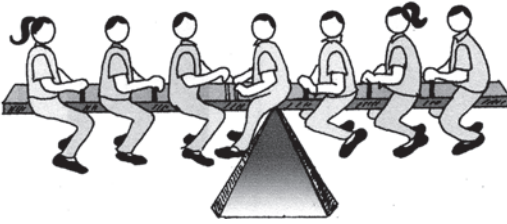
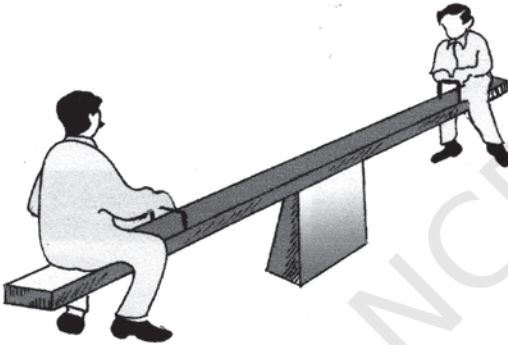
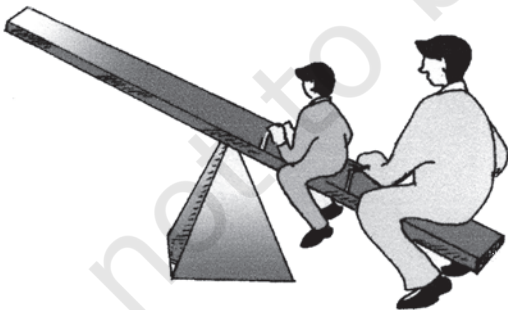
इस अध्याय का अध्ययन करने के बाद आप इस योग्य होंगे कि:

- *सहसंबंध का अर्थ समझ सकें;*
- *दो चरों के बीच संबंध के स्वरूप को समझ सकें;*
- *सहसंबंध के विभिन्न मापों का परिकलन कर सकें;*
- *संबंध की कोटि और दिशा का विश्लेषण कर सकें।*

1. प्रस्तावना

पिछले अध्याय में आपने सीखा कि आँकड़ों के समूह तथा सर्वसम चरों में परिवर्तनों का संक्षिप्त माप कैसे प्राप्त किया जाए। अब आप यह सीखेंगे कि दो चरों के बीच के संबंध का परीक्षण कैसे करें।

जैसे-जैसे गर्मी में तापमान बढ़ता है, पर्वतीय स्थलों पर सैलानियों की भीड़ बढ़ने लगती है। आइसक्रीम की बिक्री तेजी से बढ़ने लगती है। इस प्रकार, तापमान का संबंध सैलानियों की संख्या एवं आइसक्रीम की बिक्री से हो जाता है। ठीक इसी प्रकार, जब स्थानीय मंडी में टमाटर की पूर्ति बढ़ जाती है, तो उसकी कीमत कम हो जाती है। जब स्थानीय फसल तैयार होकर बाजार में पहुँचने लगती है तो टमाटरों की कीमत सामान्य पहुँच के बाहर की 40 रु प्रति किलो से घटकर 4 रु प्रति किलो या और भी कम हो जाती है। अत: पूर्ति का संबंध कीमत से रहता है। सहसंबंध का विश्लेषण ऐसे संबंधों के क्रमबद्ध परीक्षण का एक साधन है। यह निम्नलिखित प्रश्नों के समाधान करता है:

- क्या दो चरों का आपस में कोई संबंध है?

- यदि एक चर का मान बदलता है तो क्या दूसरे का मान भी बदल जाता है?

- क्या दोनों चरों में समान दिशा में परिवर्तन होता है?

- उनका यह संबंध कितना घनिष्ठ (पक्का) है?

2. संबंधों के प्रकार

आइए, पहले विभिन्न प्रकार के संबंधों पर विचार करें। माँगी गई मात्रा तथा किसी वस्तु की कीमत में परिवर्तन का संबंध माँग के सिद्धांत का अभिन्न अंग है। इसके बारे में आप विस्तार से कक्षा XII में पढ़ेंगे। कृषि उत्पादकता की कमी का संबंध बारिश की कमी से रहता है। संबंधों के इस प्रकार के उदाहरणों को कारण और परिणाम के रूप में समझा जा सकता है। अन्य उदाहरण संयोग मात्र हो सकते हैं। किसी पक्षी-विहार में प्रवासी पक्षियों के आने के साथ उस क्षेत्र में जन्म-दरों के संबंध को कारण-परिणाम संबंध का नाम नहीं दिया जा सकता। ऐसे संबंध संयोग-मात्र हैं। आपके जूते की माप और आपकी जेब में पैसों का संबंध भी संयोग का ही एक उदाहरण है, यदि इनके बीच कोई संबंध हो भी, तो उसकी व्याख्या करना कठिन होता है।

एक अन्य उदाहरण में, दो चरों पर तीसरे चर के प्रभाव से, दोनों चरों के बीच के संबंध प्रभावित हो सकते हैं। आइसक्रीम की बिक्री में तेजी डूबकर मरने वालों की संख्या से जोड़ी जा सकती है, यद्यपि मरने वाले आइसक्रीम खाकर नहीं डूबे थे। तापमान के बढ़ने के कारण ही आइसक्रीम की बिक्री में तेजी आती है। साथ ही, गर्मी से राहत पाने के लिए लोग अधिक संख्या में तरणतालों में जाने लगते हैं। संभवतः डूब कर मरने वालों की संख्या इसी कारण बढ़ गई हो। इस प्रकार, आइसक्रीम की बढ़ती हुई बिक्री और डूबने से मरने वालों की संख्या के बीच उच्च सहसंबंध का कारण तापमान है।

सहसंबंध किसका मापन करता है?

सहसंबंध चरों के बीच संबंधों की गहनता एवं दिशा का अध्ययन एवं मापन करता है। सहसंबंध सह-प्रसरण का मापन करता है न कि कार्य-कारण संबंध का।

सहसंबंध को कार्य-कारण संबंध के रूप में नहीं समझा जाना चाहिए। दो चरों x और y के बीच सहसंबंध की उपस्थिति का अर्थ है कि जब एक चर का मान किसी दिशा में बदलता है तो दूसरे चर का मान या तो उसी दिशा में बदलता है (अर्थात् (धनात्मक परिवर्तन) या फिर विपरीत दिशा में (अर्थात् ऋणात्मक परिवर्तन)। परंतु, यह एक निश्चित ढंग से होता है। इसे आसानी से समझने के लिए, यहाँ हम मान लें कि सहसंबंध, यदि है, तो रेखीय है, अर्थात दो चरों की सापेक्ष गति को ग्राफ पेपर पर एक सीधी रेखा द्वारा दिखाया जा सकता है।

सहसंबंध के प्रकार

सहसंबंध को आमतौर पर धनात्मक या ऋणात्मक सहसंबंध के रूप में वर्गीकृत किया जा सकता है। जब चरों की गति एक ही दिशा में एक साथ होती है तो सहसंबंध को धनात्मक कहा जाता है। जब आय बढ़ती है तो उपभोग में भी वृद्धि होती है। अब आय में कमी होती है तो उपभोग भी कम हो जाता है। आइसक्रीम की बिक्री तथा तापमान दोनों एक ही दिशा में गतिमान हैं। जब चर विपरीत दिशा में गतिमान हों तो सहसंबंध ऋणात्मक कहलाता है। जब सेबों की कीमत में गिरावट आती हैं तो उनकी माँग बढ़ जाती है और जब कीमत बढ़ती है तो माँग कम हो जाती हैं। जब आप पढ़ाई में अधिक समय लगाते हैं तो आपके अनुत्तीर्ण होने की संभावना कम हो जाती है और जब पढ़ाई में कम समय लगाते हैं तो अनुत्तीर्ण होने की संभावना बढ़ जाती है। ये ऋणात्मक सहसंबंध के उदाहरण हैं। यहाँ चरों की गति विपरीत दिशाओं में होती है।

3. सहसंबंध को मापने की प्रविधियाँ

सहसंबंध को मापने के लिए ये महत्वपूर्ण सांख्यिकीय उपकरण हैं: प्रकीर्ण आरेख, कार्ल पियरसन का सहसंबंध गुणांक तथा स्पीयरमैन का कोटि सहसंबंध।

प्रकीर्ण आरेख साहचर्य के स्वरूप को कोई विशिष्ट संख्यात्मक मान दिए बिना दृश्य रूप में प्रस्तुत करता है। कार्ल पियरसन का सहसंबंध-गुणांक दो चरों के बीच के रेखीय संबंधों का संख्यात्मक मापन करता है। संबंध को तब रेखीय कहा जाता है, जब इसे एक सीधी रेखा द्वारा प्रस्तुत किया जा सके। स्पीयरमैन का सहसंबंध गुणांक व्यष्टिगत मदों के बीच उनके गुणों के आधार पर निर्धारित कोटियों के द्वारा रेखीय सहसंबंध को मापा जाता है। गुण वे चर हैं, जिनका संख्यात्मक मापन संभव नहीं जैसे लोगों का बौद्धिक स्तर, शारीरिक रूप-रंग तथा ईमानदारी आदि।

प्रकीर्ण आरेख (Scatter Diagram)

प्रकीर्ण आरेख, किसी संख्यात्मक मान के बिना, संबंधों के स्वरूप की जाँच दृश्य रूप में प्रस्तुत करने की एक उपयोगी प्रविधि है। इस प्रविधि में, दो चरों के मान को ग्राफ पेपर पर बिंदुओं के रूप में आलेखित किया जाता है। प्रकीर्ण आरेख के द्वारा संबंधों के स्वरूप को काफी सही रूप में जाना जा सकता है। प्रकीर्ण आरेख में प्रकीर्ण बिंदुओं के सामीप्य की कोटि और उनकी व्यापक दिशा के आधार पर उनके आपसी संबंधों की जानकारी प्राप्त की जा सकती है। यदि सभी बिंदु एक ही रेखा पर होते हैं तो सहसंबंध परिपूर्ण होता है एवं एक (1) के बराबर होता है। यदि प्रकीर्ण बिंदु सरल रेखा के चारों तरफ फैले हुए होते हैं तो सहसंबंध निम्न माना जाता है। सहसंबंध को तब रेखीय कहा जाता है जब प्रकीर्ण बिंदु एक रेखा पर हों या रेखा के निकट हों।

प्रकीर्ण आरेख, आरेख 7.1 से 7.5 तक दिखाए गए हैं। ये हमेशा चरों के बीच के संबंधों के बारे में जानकारी देते हैं। आरेख 7.1 में प्रकीर्णन ऊपर की ओर बढ़ती हुई रेखा के आस-पास दिखाया गया है, जो एक ही दिशा में चरों के गतिमान होने का संकेत देता है। जब X बढ़ता है तो Y भी बढ़ता है, जो

धनात्मक सहसंबंध दर्शाता है। आरेख 7.2 में सारे बिंदु नीचे की ओर ढलती रेखा के आस-पास बिखरे हुए हैं। इस बार चर विपरीत दिशा में आगे बढ़ रहे हैं। जब X बढ़ता है तो Y घटता है और Y के बढ़ने पर X घटता है। यह ऋणात्मक सहसंबंध दर्शाता है। चित्र 7.3 में न तो ऐसी ऊपर उठती रेखा है और न नीचे गिरती हुई रेखा, जिनके आसपास ये बिंदु फैले हों। यह सहसंबंध न होने का उदाहरण है। आरेख 7.4 तथा 7.5 में ये बिंदु न तो ऊपर उठती रेखा के चारों ओर फैले दिखाई देते है और न नीचे गिरती रेखा के चारों ओर। ये बिंदु स्वयं रेखाओं पर ही स्थित हैं। इन्हें क्रमशः पूर्ण धनात्मक सहसंबंध तथा पूर्ण ऋणात्मक सहसंबंध कहा जाता है। प्रकीर्ण आरेख का सावधानीपूर्वक प्रेक्षण करने से हमें संबंधों की गहनता एवं स्वरूप की जानकारी प्राप्त होती है।

क्रियात्मक गतिविधि

- अपनी कक्षा के छात्रों के कद, वजन तथा उनके द्वारा दसवीं कक्षा के दो विषयों में प्राप्त अंकों के आँकड़े संगृहीत करें। इनमें से एक बार में दो चरों को लेकर उनका प्रकीर्ण आरेख बनाएँ। आप उनमें किस प्रकार का सहसंबंध देखते हैं?

कार्ल पियरसन का सहसंबंध गुणांक (Karl Pearson's Coefficient of Correlation)

इसे गुणन आधूर्ण सहसंबंध (Product Moment Correlation) तथा सरल सहसंबंध गुणांक के नामों से भी जाना जाता है। यह दो चरों X एवं Y के बीच रेखीय संबंधों के सही संख्यात्मक मान की कोटि दर्शाता है।

यह ध्यान देना महत्वपूर्ण है कि कार्ल पीयरसन के सहसंबंध गुणांक को तभी उपयोग में लाना चाहिए जब चरों के बीच रेखीय संबंध हो। जब X और Y के बीच गैर-रेखीय संबंध हो। जब X और Y के बीच गैर-रेखीय संबंध होता है तो कार्ल पीयरसन सहसंबंध की गणना भ्रामक हो सकती है। अतः यदि सही संबंध रेखीय प्रकार का है, जैसा कि चित्र 7.1, 7.2, 7.4 तथा 7.5 के प्रकीर्ण आरेखों द्वारा दर्शाया गया है, तो कार्ल पीयरसन के सहसंबंध का आगणन किया जाना चाहिए और तब यह हमको दो चरों के बीच संबंधों की गहनता को बताएगा। परंतु, यदि सही संबंध इस प्रकार का है जैसा कि चित्र 7.6 अथवा 7.7 के प्रकीर्ण आरेखों द्वारा दिखाया गया है, तो इसका अर्थ है कि X तथा Y के बीच गैर-रेखीय संबंध है तथा हमको कार्ल पीयरसन के सहसंबंध गुणांक का उपयोग करने का प्रयास नहीं करना चाहिए।

अतः यह उचित है कि पहले चरों के बीच संबंध के प्रकीर्ण चित्र की कार्ल पीयरसन के सहसंबंध गुणांक की गणना से पूर्व, जाँच की जाए।

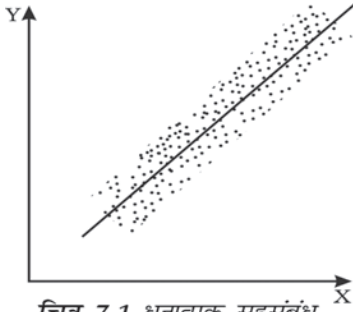
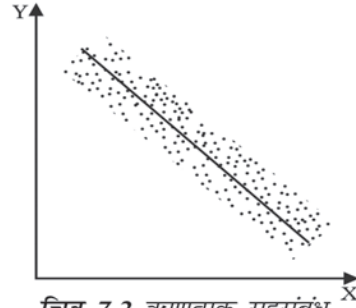
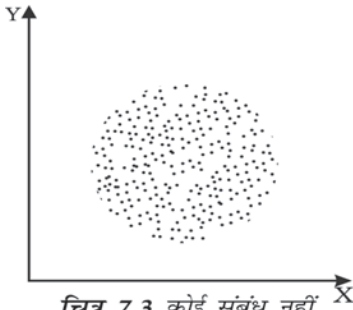
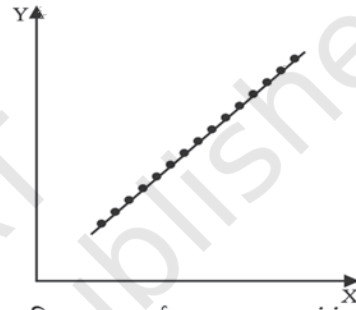
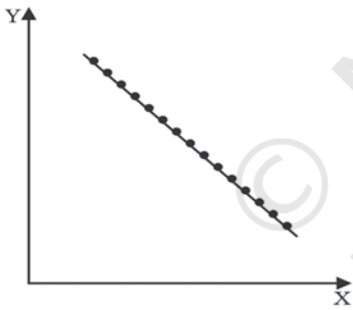
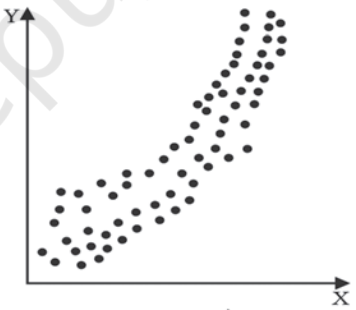
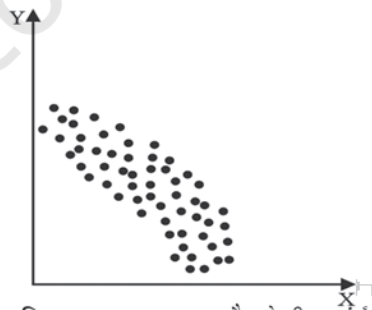
मान लें कि $X_1, X_2, \ldots, X_N$ आदि X के N मान हैं तथा $Y_1, Y_2, \ldots, Y_N$ Y के संगत मान हैं। आगे की प्रस्तुतियों में सरलता की दृष्टि से इकाइयों को दर्शाने वाले पादांकों को छोड़ दिया गया है। X तथा Y के समांतर माध्य को इस प्रकार परिभाषित किया गया है:

$$\bar{X} = \frac{\Sigma X}{N}; \qquad \bar{Y} = \frac{\Sigma Y}{N}$$

और उनके प्रसरण निम्नलिखित हैं:

$$\sigma^2{}_x = \frac{\Sigma(X - \bar{X})^2}{N} = \frac{\Sigma X^2}{N} - \bar{X}^2$$

तथा $\sigma^2{}_y = \dfrac{\Sigma(Y - \bar{Y})^2}{N} = \dfrac{\Sigma Y^2}{N} - \bar{Y}^2$

यहाँ, X एवं Y के मानक विचलन क्रमशः उनके प्रसरण के धनात्मक वर्गमूल हैं। X तथा Y के सहप्रसरण निम्नलिखित हैं:

चित्र 7.1 धनात्मक सहसंबंध

चित्र 7.2 ऋणात्मक सहसंबंध

चित्र 7.3 कोई संबंध नहीं

चित्र 7.4 पूर्ण धनात्मक सहसंबंध

चित्र 7.5 पूर्ण ऋणात्मक सहसंबंध

चित्र 7.6 धनात्मक गैर-रेखीय संबंध

चित्र 7.7 ऋणात्मक गैर-रेखीय संबंध

$$\text{Cov}(X,Y) = \frac{\Sigma(X-\bar{X})(Y-\bar{Y})}{N} = \frac{\Sigma xy}{N}$$

जहाँ $x = X-\bar{X}$ तथा $y = Y-\bar{Y}$ । ये X तथा Y के माध्य मानों से उनके i वें मान के विचलन हैं।

X और Y के बीच सहप्रसरण का चिह्न सहसंबंध गुणांक के चिह्न का निर्धारण करता है। मानक विचलन हमेशा धनात्मक होते हैं। यदि सहप्रसरण शून्य होता है, तो सहसंबंध गुणांक भी सदैव शून्य होता है। गुणन आघूर्ण सहसंबंध या कार्ल पियरसन का सहसंबंध मापन नीचे दिया जा रहा है,

$$r = \Sigma xy \Big/ N\sigma_x\sigma_y \qquad ...(1)$$

या

$$r = \frac{\Sigma(X-\bar{X})(Y-\bar{Y})}{\sqrt{\Sigma(X-\bar{X})^2}\sqrt{\Sigma(Y-\bar{Y})^2}} \qquad ...(2)$$

या

$$r = \frac{\Sigma XY - \frac{(\Sigma X)(\Sigma Y)}{N}}{\sqrt{\Sigma X^2 - \frac{(\Sigma X)^2}{N}}\sqrt{\Sigma Y^2 - \frac{(\Sigma Y)^2}{N}}} \qquad ...(3)$$

या

$$r = \frac{N\Sigma XY - (\Sigma X)(\Sigma Y)}{\sqrt{N\Sigma X^2 - (\Sigma X)^2}\sqrt{N\Sigma Y^2 - (\Sigma Y)^2}} \qquad ...(4)$$

सहसंबंध गुणांक के गुण

सहसंबंध गुणांक के गुण निम्नलिखित हैं:

- r की कोई इकाई नहीं होती। यह एक संख्या-मात्र है। इसका तात्पर्य है कि माप की इकाइयाँ r का हिस्सा नहीं हैं। उदाहरण के लिए, कद (फुटों में) तथा वजन (कि.ग्रा. में) के बीच r है 0.7।
- r का ऋणात्मक मान प्रतिलोम संबंध दर्शाता है। किसी चर में बदलाव, दूसरे चर में विपरीत दिशा में बदलाव के साथ संबंद्ध रहता है। जब एक वस्तु की कीमत बढ़ती है तो उसकी माँग घट जाती है। जब ब्याज दर बढ़ती है तो निधियों (ब्याज पर ली जाने वाली धन-राशियाँ) की माँग घट जाती है। ऐसा इसलिए होता है क्योंकि निधियाँ महँगी हो जाती हैं।

- यदि r धनात्मक होता है तो दोनों चर एक ही दिशा में गतिमान होते हैं। जब चाय के स्थानापन्न के रूप में कॉफी के दाम बढ़ते हैं, तो चाय की माँग भी बढ़ जाती है। सिंचाई व्यवस्था के सुधार का संबंध फसलों की अधिक पैदावार से रहता है। जब तापमान में वृद्धि होती है, तो आइसक्रीम की बिक्री बढ़ जाती है।
- सहसंबंध गुणांक का मान –1 तथा +1 के बीच स्थित होता है $-1 \leq r \leq +1$। यदि किसी भी अभ्यास में r का मान इस परास के बाहर होता है तो इससे परिकलन में त्रुटि का संकेत मिलता है।
- 'r' परिमाण, उद्गम और पैमाने के परिवर्तन से अप्रभावित होता है। यदि हमें दो चर X तथा Y दिए गए हों तो दो नए चरों को इस प्रकार परिभाषित किया जा सकता है–

$$U = \frac{X-A}{B} \quad ; \qquad V = \frac{Y-C}{D}$$

यहाँ पर A तथा C क्रमशः X तथा Y के कल्पित मान हैं। B तथा D समापवर्तक हैं और इनका समान उद्‌गम है।

अतः

$r_{xy} = r_{uv}$

अति सरल प्रकार से, सहसंबंध गुणांक की गणना में, पद विचलन पद्धति की भाँति, इस गुण का उपयोग किया जाता है।

- $r = 0$, तो इसका अर्थ है कि दो चरों में सह संबंध नहीं है। उनके बीच कोई रेखीय संबंध नहीं है। वैसे, अन्य प्रकार के संबंध हो सकते हैं।
- $r = 1$ अथवा $r = -1$, तो इसका अर्थ है कि सहसंबंध पूर्ण है और चरों के बीच सटीक रेखीय संबंध है।
- r के मान का होना, घनिष्ठ रेखीय संबंध को इंगित करता है। इसके मान को उच्च तब कहा जाता है जब यह +1 अथवा –1 के निकट होता है।
- r का निम्न मान (शून्य के निकट), मंद रेखीय संबंध को इंगित करता है, परंतु गैर-रेखीय संबंध पाया जा सकता है।

हमने पहले अध्याय में चर्चा की है कि सांख्यिकीय विधियाँ व्यवहार बुद्धि का स्थानापन्न नहीं हैं। एक अन्य उदाहरण लेते हैं, जो सहसंबंध के परिकलन और व्याख्या से पहले आँकड़ों की विशेषताओं को समझने की आवश्यकता पर प्रकाश डालता है। कुछ गाँवों में महामारी फैलती है और सरकार प्रभावित गाँवों में डॉक्टरों का दल भेजती है। गाँव में होने वाली मौतों की संख्या तथा भेजे गए डॉक्टरों की संख्या के बीच धनात्मक सहसंबंध पाया गया। (अर्थात् डॉक्टरों की संख्या बढ़ने से मौतें बढ़ गई)। सामान्यतः डॉक्टरों द्वारा उपलब्ध कराई जानेवाली सेवाओं के परिणामस्वरूप मृत्यु दर में कमी की आशा की जाती है, अर्थात् इनके बीच ऋणात्मक सहसंबंध होता है। यदि ऐसा नहीं हुआ, तो इसके पीछे अन्य कारण रहे होंगे। आँकड़े, संभवतः, किसी अवधि-विशेष से संबंधित होंगे या फिर, दर्ज की गई मृत्यु दर संभवतः ऐसे व्यक्तियों के बारे में हो सकती है जिनकी दशा बहुत बिगड़ चुकी थी। साथ ही, किसी भी क्षेत्र में डाक्टरों की उपस्थिति का सुपरिणाम कुछ समय बीतने के बाद ही दिखाई देता है। यह भी संभव है कि दर्ज की गई मौतें महामारी के कारण हुई ही न हों। जैसे, सुनामी ने अचानक किसी देश में अपना भयंकर रूप दिखाया हो और मृत्यु-दर बढ़ गई हो।

आइए, किसानों द्वारा विद्यालय में बिताए गए वर्षों तथा प्रति एकड़ वार्षिक उपज के बीच के संबंध के परीक्षण के द्वारा r के परिकलन को सोदाहरण स्पष्ट करें:

उदाहरण 1

किसानों द्वारा विद्यालय में बिताए गए वर्ष	*प्रति एकड़ वार्षिक उपज ('000 रु में)*
0	4
2	4
4	6
6	10
8	10
10	8
12	7

सूत्र 1 के लिए $\sum xy, \sigma_x, \sigma_y$ के मानों की आवश्यकता है। सारणी 7.1 के द्वारा हम इन मान को प्राप्त कर सकते हैं।

$$\sum xy = 42,$$

$$\sigma_x = \sqrt{\frac{\sum(X-\bar{X})^2}{N}} = \sqrt{\frac{112}{7}},$$

$$\sigma_y = \sqrt{\frac{\Sigma(Y-\bar{Y})^2}{N}} = \sqrt{\frac{38}{7}}$$

इन मानों को सूत्र 1 में प्रतिस्थापित करने पर,

$$r = \frac{42}{7\sqrt{\frac{112}{7}}\sqrt{\frac{38}{7}}} = 0.644$$

सूत्र 2 के द्वारा भी इन्हीं मानों को प्राप्त किया जा सकता है,

$$r = \frac{\Sigma(X-\bar{X})(Y-\bar{Y})}{\sqrt{\Sigma(X-\bar{X})^2}\sqrt{\Sigma(Y-\bar{Y})^2}} \quad ...(2)$$

$$r = \frac{42}{\sqrt{112}\ \sqrt{38}} = 0.644$$

इस प्रकार, हमने देखा कि किसानों की शिक्षा के वर्ष तथा प्रति एकड़ उपज के बीच धनात्मक सहसंबंध है। साथ ही r का मान भी अधिक है। इससे पता चलता है कि किसान जितने अधिक वर्षों तक शिक्षा ग्रहण करेंगे, प्रति एकड़ उपज उतनी ही अधिक होगी। इससे किसानों के लिए शिक्षा के महत्त्व पर प्रकाश पड़ता है।

सूत्र (3) का प्रयोग करने पर

$$r = \frac{\Sigma XY - \frac{(\Sigma X)(\Sigma Y)}{N}}{\sqrt{\Sigma X^2 - \frac{(\Sigma X)^2}{N}}\sqrt{\Sigma Y^2 - \frac{(\Sigma Y)^2}{N}}} \quad ...(3)$$

इस सूत्र के प्रयोग के लिए हमें निम्नलिखित व्यंजकों का परिकलन करना होगा,

$\Sigma XY, \Sigma X^2, \Sigma Y^2$.

अब r का मूल्य जानने के लिए सूत्र (3) का प्रयोग करें।

आइए, अब r के मान की विभिन्न व्याख्याओं की जानकारी लें। मान लें कि अंग्रेजी तथा सांख्यिकी इन दोनों विषयों के प्राप्तांकों के बीच सहसंबंध 0.1 है। इसका अर्थ है कि इन दोनों विषयों में प्राप्त किए गए अंकों में धनात्मक सहसंबंध है एवं सहसंबंध की प्रबलता कमजोर है। अंग्रेजी में अधिक अंक प्राप्त करने वाले छात्र सांख्यिकी में अपेक्षाकृत कम अंक प्राप्त कर सकते हैं। यदि r का मान 0.9 होता, तो अंग्रेजी में अधिक प्राप्तांक वाले विद्यार्थियों ने निश्चित रूप से सांख्यिकी में अधिक अंक प्राप्त किए होते।

सारणी 7.1

किसानों की शिक्षा के वर्ष एवं प्रति एकड़ पैदावार के बीच r का परिकलन

शिक्षा के वर्ष (x)	$(X-\bar{X})$	$(X-\bar{X})^2$	प्रति एकड़ वार्षिक पैदावार '000 रु (Y)	$(Y-\bar{Y})$	$(Y-\bar{Y})^2$	$(X-\bar{X})(Y-\bar{Y})$
0	−6	36	4	−3	9	18
2	−4	16	4	−3	9	12
4	−2	4	6	−1	1	2
6	0	0	10	3	9	0
8	2	4	10	3	9	6
10	4	16	8	1	1	4
12	6	36	7	0	0	0
$\Sigma X=42$	$\Sigma(X-\bar{X})^2=112$		$\Sigma Y=49$	$\Sigma(Y-\bar{Y})^2=38$		$\Sigma(X-\bar{X})(Y-\bar{Y})=42$

ऋणात्मक सहसंबंध के एक उदाहरण के रूप में स्थानीय मंडी में सब्जियों के आगमन के साथ उनकी कीमत के संबंध को लिया जा सकता है। यदि r = −0.9 होता, तो स्थानीय मंडी में सब्जियों की पूर्ति बढ़ने के साथ इनकी कीमत कम होगी। यदि r = −0.1 होता तो भी सब्जियों की अधिक पूर्ति के साथ इनकी कीमतें कम तो होतीं, परंतु उतनी कम नहीं जितनी तब थीं, जब r = −0.9 था। कीमत में किस हद तक गिरावट होगी इसका संबंध r के निरपेक्ष मान के साथ है। यदि r = 0 होगा, तो बाजार में पूर्ति के काफी बढ़ने पर भी, कीमत में कोई कमी नहीं होती। ऐसी भी संभावना है कि पूर्ति के बढ़ने पर कुशल परिवहन तंत्र की सहायता से इन्हें अन्य बाजारों में ले जाया गया हो।

क्रियात्मक गतिविधि

- निम्नलिखित सारणी को देखें। वर्तमान कीमत पर राष्ट्रीय आय में वार्षिक वृद्धि तथा (सकल घरेलू उत्पाद के प्रतिशत के रूप में) सकल घरेलू बचत के बीच r का परिकलन कीजिए।

सारणी 7.2

वर्ष	*राष्ट्रीय आय की वार्षिक वृद्धि*	*सकल घरेलू बचत GDP के प्रतिशत के रूप में*
1992−93	14	24
1993−94	17	23
1994−95	18	26
1995−96	17	27
1996−97	16	25
1997−98	12	25
1998−99	16	23
1999−00	11	25
2000−01	8	24
2001−02	10	23

स्रोत: *आर्थिक सर्वेक्षण, (2004−05) पृष्ठ 8, 9*

सहसंबंध गुणांक के परिकलन में पद-विचलन विधि

जब चरों के मान ऊँचे हों, तो परिकलन की समस्या को r के एक गुण के प्रयोग द्वारा कम किया जा सकता है। यह गुण है कि r 'उद्गम परिवर्तन' तथा 'स्केल परिवर्तन' से प्रभावित नहीं होता है। इसे पद विचलन विधि के रूप में भी जाना जाता है। इसके अंतर्गत X एवं Y चरों को निम्नलिखित पद विचलन विधि से परिवर्तित किया जा सकता है:

$$U = \frac{X - A}{B}; V = \frac{Y - C}{D}$$

यहाँ A तथा B कल्पित माध्य हैं तथा h एवं k समापवर्तक हैं एवं एक ही चिह्न के हैं।

अत: $r_{UV} = r_{XY}$

इसे कीमत सूचकांक तथा धन की पूर्ति के बीच सहसंबंध के विश्लेषण की प्रक्रिया के द्वारा समझा जा सकता है।

उदाहरण 2

कीमत सूचकांक (X)	120	150	190	220	230
धन की पूर्ति (करोड़ रु में) (Y)	1800	2000	2500	2700	3000

पद विचलन विधि का प्रयोग करते हुए, सरलीकरणों को निम्नलिखित विधि द्वारा दिखाया गया है:

A = 100; h = 10; B = 1700 एवं k = 100

चरों की रूपांतरित सारणी नीचे दी गई है:

कीमत सूचकांक तथा मुद्रा की पूर्ति के बीच पद-विचलन विधि का उपयोग करते हुए r का परिकलन:

सारणी 7.3

U $\left(\frac{X-100}{10}\right)$	V $\left(\frac{Y-1700}{100}\right)$	U^2	V^2	UV
2	1	4	1	2
5	3	25	9	15
9	8	81	64	72
12	10	144	100	120
13	13	169	169	169

$\Sigma U = 41;\ \Sigma V = 35;\ \Sigma U^2 = 423;$
$\Sigma V^2 = 343;\ \Sigma UV = 378$

इन मानों को सूत्र (3) में प्रतिस्थापन करने पर

$$r = \frac{\Sigma UV - \frac{(\Sigma U)(\Sigma V)}{N}}{\sqrt{\Sigma U^2 - \frac{(\Sigma U)^2}{N}}\ \sqrt{\Sigma V^2 - \frac{(\Sigma V)^2}{N}}} \qquad (3)$$

$$= \frac{378 - \frac{41 \times 35}{5}}{\sqrt{423 - \frac{(41)^2}{5}}\ \sqrt{343 - \frac{(35)^2}{5}}}$$

$= 0.98$

कीमत सूचकांक एवं मुद्रा-पूर्ति के बीच यह प्रबल धनात्मक सहसंबंध वित्तीय नीतियों के लिए महत्त्वपूर्ण आधार है। जब मुद्रा-पूर्ति बढ़ती है तब कीमत सूचकांक में भी वृद्धि होती है।

क्रियात्मक गतिविधि

- भारत की जनसंख्या एवं राष्ट्रीय आय से संबंधित आँकड़ों का उपयोग करें और पद विचलन विधि का उपयोग करते हुए उनके बीच सहसंबंध का परिकलन करें।

स्पीयरमैन का कोटि सहसंबंध (Spearman's Rank Correlation)

'स्पीयरमैन कोटि सहसंबंध' का विकास ब्रिटिश मनोवैज्ञानिक सी.ई. स्पीयरमैन द्वारा किया गया था। इसका उपयोग निम्न परिस्थितियों में किया जाता है–

1. कल्पना कीजिए कि हमें किसी दूर-दराज़ के गाँव में जहाँ न कोई मापदंड उपलब्ध है और न कोई वज़न मापने की कोई मशीन, छात्रों की लंबाई और वज़न के बीच, सहसंबंध का आकलन करना है। ऐसी स्थिति में हम लंबाई अथवा वज़न का माप नहीं कर सकते, परंतु हम छात्रों को उनकी लंबाई और वज़न के अनुसार निश्चित रूप से कोटिबद्ध कर सकते हैं और फिर इन कोटियों को स्पीयरमैन के सहसंबंध की गणना में उपयोग किया जा सकता है।
2. कल्पना कीजिए कि हमें, निष्पक्षता, ईमानदारी अथवा सौंदर्य का अध्ययन करना है। हम इनका उसी प्रकार माप नहीं कर सकते, जिस प्रकार आय, भार अथवा लंबाई का। अधिक से अधिक, इन चीज़ों का सापेक्ष माप किया जा सकता है। उदाहरण के लिए, हम लोगों को सौंदर्य के आधार पर कोटिबद्ध कर सकते हैं। कुछ लोग यह बहस कर सकते हैं कि ऐसा करना संभव नहीं है, क्योंकि सौंदर्य मापने के मापदंड और कसौटियाँ, व्यक्ति से व्यक्ति तथा संस्कृति से संस्कृति भिन्न हो सकती है। यदि हमें दो चरों के बीच, जिनमें कम से कम एक उपरोक्त प्रकार का है, तो स्पीयरमैन के सहसंबंध गुणांक का उपयोग किया जाएगा।
3. स्पीयरमैन के कोटि सहसंबंध का उन स्थितियों में भी उपयोग किया जा सकता है, जिनमें संबंध की दिशा तो स्पष्ट है, लेकिन वह गैर-रेखीय है, जैसा कि चित्र 7.6 तथा 7.7 के प्रकीर्ण चित्रों द्वारा प्रदर्शित किया गया है।

4. स्पीयरमैन का सहसंबंध गुणांक चरम मूल्यों से प्रभावित नहीं होता। इस दृष्टि से यह कार्ल पीयरसन के सहसंबंध गुणांक से उत्तम है। अतः समंकों में यदि कुछ चरम मूल्य हैं, तो स्पीयरमैन के सहसंबंध गुणांक का उपयोग अति लाभप्रद होता है।

कोटि सहसंबंध गुणांक तथा सरल सहसंबंध गुणांक की व्याख्या समान रूप से की जाती है। इसका सूत्र सरल सहसंबंध गुणांक से प्राप्त किया गया है जहाँ व्यष्टिगत मानों को कोटियों द्वारा प्रतिस्थापित किया जाता है। इन कोटियों का प्रयोग सहसंबंध के परिकलन के लिए किया जाता है। यह गुणांक इन इकाइयों के लिए निर्धारित कोटियों के बीच रेखीय संबंध को मापता है, न कि उनके मानों के बीच। स्पीयरमैन का कोटि सहसंबंध निम्नलिखित सूत्र द्वारा प्राप्त करते हैं:

$$r_s = 1 - \frac{6\Sigma D^2}{n^3 - n} \qquad ...(4)$$

यहाँ 'n' प्रेक्षणों की संख्या है तथा D किसी चर के लिए निर्धारित कोटियों का, किसी अन्य चर के लिए निर्धारित कोटि से, विचलन दर्शाता है।

सरल सहसंबंध गुणांक के सभी गुण यहाँ लागू किए जा सकते हैं। पियरसन सहसंबंध गुणांक की भाँति यह भी +1 तथा –1 के बीच स्थित होता है। हालाँकि, सामान्य तौर पर यह सामान्य विधि की तरह यथातथ नहीं होता है। इसका कारण यह है कि आँकड़ों से संबद्ध सभी सूचनाओं का उपयोग नहीं होता है।

प्रथम अंतर क्रमिक मानों में अंतर होता है। श्रृंखला में मदों के मानों के वे प्रथम अंतर जो उनके परिमाण के अनुसार क्रम में व्यवस्थित किए जाते हैं, आमतौर पर कभी स्थिर नहीं होते। सामान्यतः आँकड़ा-गुच्छ केंद्रीय मानों के आस पास सरणी के मध्य में थोड़े बहुत अंतर पर एकत्र होता है।

यदि प्रथम अंतर स्थिर होते, तब r और r_k समान परिमाण देते। सामान्यतः r_k का मान r से कम या इसके बराबर होता है।

कोटि सहसंबंध का परिकलन

1. जब कोटियाँ दी गई हों।
2. जब कोटियाँ नहीं दी गई हों। उन्हें आँकड़ों से प्राप्त किया जाना हो।
3. जब कोटियों की पुनरावृत्ति की गई हो।

स्थिति 1: जब कोटियाँ दी गई हों

उदाहरण 3

किसी सौंदर्य प्रतियोगिता में तीन निर्णायकों द्वारा पाँच लोगों का मूल्यांकन किया जाता है। हमें ज्ञात करना है कि सौंदर्य-बोध के प्रति किन दो निर्णायकों का दृष्टिकोण सर्वाधिक समान है।

	प्रतियोगी				
निर्णायक	1	2	3	4	5
क	1	2	3	4	5
ख	2	4	1	5	3
ग	1	3	5	2	4

यहाँ पर निर्णायकों के तीन जोड़े हैं, अतः कोटि सहसंबंध का परिकलन तीन बार किया जायगा। यहाँ सूत्र (4) का प्रयोग करना चाहिए,

$$r_s = 1 - \frac{6\Sigma D^2}{n^3 - n} \qquad ...(4)$$

निर्णायकों क और ख के बीच कोटि-सहसंबंध नीचे परिकलित किया गया है:

क	*ख*	*ग*	*ग²*
1	2	−1	1
2	4	−2	4
3	1	2	4
4	5	−1	1
5	3	2	4
योग			*14*

सूत्र (4) में इन मानों को प्रतिस्थापित करने पर

$$r_s = 1 - \frac{6\Sigma D^2}{n^3 - n} \qquad ...(4)$$

$$= 1 - \frac{6 \times 14}{5^3 - 5} = 1 - \frac{84}{120} = 1 - 0.7 = 0.3$$

निर्णायकों (क) और (ग) के बीच कोटि सहसंबंध निम्नवत् परिकलित किया गया है:

क	*ख*	*ग*	*ग²*
1	1	0	0
2	3	−1	1
3	5	−2	4
4	2	2	4
5	4	1	1
योग			*10*

सूत्र (4) में इन मानों को प्रतिस्थापित करने पर कोटि सहसंबंध 0.5 होता है। ठीक इसी प्रकार से निर्णायकों 'ख' और 'ग' के बीच कोटि सहसंबंध 0.9 है। अत: निर्णायकों 'क' और 'ग' के सौंदर्य बोध निकटतम हैं। निर्णायक 'ख' और 'ग' की रुचियाँ काफी भिन्न है।

स्थिति 2: जब कोटियाँ नहीं दी गई हों

उदाहरण 4

यहाँ पर 5 छात्रों द्वारा अर्थशास्त्र एवं सांख्यिकी विषयों में प्राप्त अंकों का प्रतिशत दिया गया है। अब कोटियों का निर्धारण करना है और कोटि सह-संबंध का परिकलन करना है।

छात्र	*सांख्यिकी में प्राप्तांक (X)*	*अर्थशास्त्र में प्राप्तांक (Y)*
क	85	60
ख	60	48
ग	55	49
घ	65	50
ड़	75	55

छात्र	*सांख्यिकी में कोटियाँ (R_x)*	*अर्थशास्त्र में कोटियाँ (R_Y)*
क	1	1
ख	4	5
ग	5	4
घ	3	3
ङ	2	2

एक बार जब कोटियाँ देने का क्रम जब पूरा हो जाए तो कोटि सहसंबंध के परिकलन के लिए सूत्र (4) का प्रयोग किया जाता है।

स्थिति 3: जब कोटियों को दोहराया गया हो

उदाहरण 5

X तथा Y के मान नीचे दिए गये हैं:

(X)	*(Y)*
1200	75
1150	65
1000	50
990	100
800	90
780	85
760	90
750	40
730	50
700	60
620	50
600	75

कोटि सहसंबंध के परिकलन के लिए मानों की कोटियाँ निर्धारित की जाती हैं। दोहराए गए मदों के लिए समान कोटियाँ दी जाती हैं। समान कोटि उन कोटियों का माध्य है जिन्हें वे मद तब धारण करते हैं, जब उनमें एक दूसरे से भिन्नता होती। अगले मद के लिए वह कोटि निर्धारित की जायेगी जो पहले दी गई कोटि के बाद होगी।

यहाँ नौवीं, दसवीं तथा ग्यारहवीं कोटियों का मान 50 है। अत: इन तीनों को औसत कोटि अर्थात 10 दी गई है।

कोटि X	*कोटि Y*	*कोटि क्रम में विचलन*	*D^2*
1	5.5	−4.5	20.25
2	7	−5	25.00
3	10	−7	49.00
4	1	3	9.00
5	2.5	2.5	6.25
6	4	2	4.00
7	2.5	4.5	20.25
8	12	−4	16.00
9	10	−1	1.00
10	8	2	4.00
11	10	1	1.00
12	5.5	6.5	42.25
योग			*198.00*

जब कोटियों को दोहराया जाता है तो स्पीयरमैन कोटि सहसंबंध के गुणांक का सूत्र इस प्रकार है–

$$r_s = 1 - \frac{6\left[\Sigma D^2 + \frac{(m^3{}_1 - m_1)}{12} + \frac{(m^3{}_2 - m_2)}{12} + \ldots\right]}{n(n^2 - 1)}$$

यहाँ $m_1, m_2, \ldots,$ कोटियों की पुनरावृत्त संख्याएँ हैं और $\frac{m^3{}_1 - m_1}{12} \ldots,$ उनके संगत संशोधन गुणक हैं। इस विवरण के लिए आवश्यक सुधार इस प्रकार है:

$$\frac{3^3-3}{12} + \frac{2^3-2}{12} = \frac{30}{12} = 2.5$$

इन व्यंजकों के मानों को प्रतिस्थापित करने पर,

$$rs = 1 - \frac{6(198 + 2.5)}{12^3 - 12} = (1 - 0.70) = 0.30$$

इस प्रकार यहाँ पर X और Y के बीच धनात्मक कोटि सहसंबंध है। X तथा Y दोनों एक ही दिशा में गतिमान हैं। हालाँकि इनके संबंध को सुदृढ़ नहीं कहा जा सकता।

क्रियात्मक गतिविधि

- अपनी कक्षा के 10 छात्रों द्वारा नवीं और दसवीं की परीक्षाओं में प्राप्त किए अंकों के आँकड़े संगृहीत करें। उनके बीच कोटि सहसंबंध गुणांक का परिकलन करें। यदि आपके आँकड़ों में पुनरावर्तन हो, तो दोहराई गई कोटियों वाले आँकड़ों का संग्रह करके इस अभ्यास को पुनः दोहराएँ।

  ऐसी कौन सी स्थितियाँ हैं, जिनमें कोटि सहसंबंध गुणांक को सरल सह संबंध गुणांक की तुलना में प्राथमिकता दी जाती है। यदि आँकड़ों को सही ढंग से मापा जाय, तो क्या फिर भी आप कोटि सहसंबंध गुणांक की तुलना में सरल गुणांक को प्राथमिकता देंगे? आप किन स्थितियों में इनके चुनाव में तटस्थ रह सकते हैं? कक्षा में इन मुद्दों पर चर्चा कीजिए।

4. सारांश

हमने दो चरों के बीच संबंध, विशेषतः रेखीय संबंध , के अध्ययन के लिए कुछ प्रविधियों की चर्चा की। प्रकीर्ण आरेख संबंधों की दृश्यात्मक प्रस्तुति करता है और यह रेखीय संबंध तक ही सीमित नहीं है। कार्ल पियरसन का सहसंबंध गुणांक तथा स्पीयरमैन का कोटि-सहसंबंध चरों के बीच रेखीय संबंधों की माप हैं। जब चरों को परिशुद्ध रूप से मापना संभव न हो, तो वहाँ कोटि सहसंबंध का प्रयोग हो सकता है। लेकिन ये माप कार्य-कारण संबंध सूचित नहीं करते। जब सहसंबंधित चरों में परिवर्तन होता है, तो सहसंबंध का ज्ञान हमें चरों में परिवर्तन की दिशा तथा गहनता के बारे में बताता है।

पुनरावर्तन

- सहसंबंध विश्लेषण के अंतर्गत दो चरों के बीच के संबंधों का अध्ययन किया जाता है।
- प्रकीर्ण आरेख दो चरों के बीच संबंध के स्वरूप का दृश्य प्रस्तुतीकरण करता है।
- कार्ल पियरसन का सहसंबंध गुणांक r दो चरों के बीच केवल रेखीय संबंध को संख्यात्मक रूप से मापता है। r सदैव –1 तथा +1 के बीच स्थित रहता है।
- यदि चरों को परिशुद्धता से न मापा जा सके, तो स्पीयरमेन के कोटि सहसंबंध का उपयोग रेखीय संबंधों को संख्यात्मक रूप से मापने के लिए किया जा सकता है।
- दोहराई गई कोटियों को संशोधन गुणकों की आवश्यकता होती है।
- सहसंबंध का तात्पर्य कार्य–कारण संबंध नहीं, बल्कि केवल सहप्रसरण दर्शाना है।

अभ्यास

1. कद (फुटों में) तथा वज़न (किलोग्राम में) के बीच सहसंबंध गुणांक की इकाई है:
   (क) कि.ग्रा./फुट
   (ख) प्रतिशत
   (ग) अविद्यमान
2. सरल सहसंबंध गुणांक का परास निम्नलिखित होगा
   (क) 0 से अनंत तक
   (ख) –1 से +1 तक
   (ग) ऋणात्मक अनंत (infinity) से धनात्मक अनंत (infinity) तक
3. यदि r_{xy} धनात्मक है तो x और y के बीच का संबंध इस प्रकार का होता है:
   (क) जब y बढ़ता है तो x बढ़ता है।
   (ख) जब y घटता है तो x बढ़ता है।
   (ग) जब y बढ़ता है तो x नहीं बदलता है।
4. यदि $r_{xy} = 0$ तब चर x और y के बीच:
   (क) रेखीय संबंध होगा
   (ख) रेखीय संबंध नहीं होगा
   (ग) स्वतंत्र होगा
5. निम्नलिखित तीनों मापों में, कौन सा माप किसी भी प्रकार के संबंध की माप कर सकता है।
   (क) कार्ल पियरसन सहसंबंध गुणांक
   (ख) स्पीयरमैन का कोटि सहसंबंध
   (ग) प्रकीर्ण आरेख
6. यदि परिशुद्ध रूप से मापित आँकड़े उपलब्ध हों, तो सरल सहसंबंध गुणांक:
   (क) कोटि सहसंबंध गुणांक से अधिक सही होता है।
   (ख) कोटि सहसंबंध गुणांक से कम सही होता है।
   (ग) कोटि सहसंबंध की ही भाँति सही होता है।

7. साहचर्य के माप के लिए r को सहप्रसरण से अधिक प्राथमिकता क्यों दी जाती है?
8. क्या आँकड़ों के प्रकार के आधार पर r, –1 तथा +1 के बाहर स्थित हो सकता है?
9. क्या सहसंबंध के द्वारा कार्यकारण संबंध की जानकारी मिलती है?
10. सरल सहसंबंध गुणांक की तुलना में कोटि सहसंबंध गुणांक कब अधिक परिशुद्ध होता है?
11. क्या शून्य सहसंबंध का अर्थ स्वतंत्रता है?
12. क्या सरल सहसंबंध गुणांक किसी भी प्रकार के संबंध को माप सकता है?
13. एक सप्ताह तक अपने स्थानीय बाजार से 5 प्रकार की सब्जियों की कीमतें प्रतिदिन एकत्र करें। उनका सहसंबंध गुणांक परिकलित कीजिए। इसके परिणाम की व्याख्या कीजिए।
14. अपनी कक्षा के सहपाठियों के कद मापिए। उनसे उनके बेंच पर बैठे सहपाठी का कद पूछिए। इन दो चरों का सहसंबंध गुणांक परिकलित कीजिए और परिणाम का निर्वचन कीजिए।
15. कुछ ऐसे चरों की सूची बनाएँ जिनका परिशुद्ध मापन कठिन हो।
16. r के विभिन्न मानों +1, –1, तथा 0 की व्याख्या करें।
17. पियरसन सहसंबंध गुणांक से कोटि सहसंबंध गुणांक क्यों भिन्न होता है?
18. पिताओं (x) और उनके पुत्रों (y) के कदों का माप नीचे इंचों में दिया गया है, इन दोनों के बीच सहसंबंध गुणांक को परिकलित कीजिए

x	65	66	57	67	68	69	70	72
y	67	56	65	68	72	72	69	71

(उत्तर r = 0.603)

19. x और y के बीच सहसंबंध गुणांक को परिकलित कीजिए और उनके संबंध पर टिप्पणी कीजिए।

x	–3	–2	–1	1	2	3
y	9	4	1	1	4	9

(उत्तर r = 0)

20. x और y के बीच सहसंबंध गुणांक को परिकलित कीजिए और उनके संबंध पर टिप्पणी कीजिए।

x	1	3	4	5	7	8
y	2	6	8	10	14	16

(उत्तर r = 1)

क्रियात्मक गतिविधि

- भारत की राष्ट्रीय आय और निर्यात के कम से कम 10 प्रेक्षण लेकर, इस पाठ में बताए गए सभी सूत्रों का उपयोग करते हुए r को परिकलित कीजिए।